DOCUMENT RESUME

ED 183 370 SE 029 898

AUTHOR Corsale, Kathleen; Gitomer, Drew
TITLE Knowledge Base Effects in Children's Number Analogy Solutions.
PUB DATE Nov 79
NOTE 15p.: Paper presented at the meeting of the Psychonomic Society (Phoenix, AZ, November 1979); Not available in hard copy due to copyright restrictions

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Academic Aptitude; Elementary Secondary Education; *Individual Differences; Intelligence Differences; *Knowledge Level; *Mathematics Education; Number Concepts; *Problem Solving; *Research

ABSTRACT

Developmental and individual differences in mathematical aptitude were investigated as a function of knowledge structure and processing variables. Results indicated the relative importance of knowledge structure and strategy skills in aptitude test performance. Protocol data further elaborated the interrelationships between knowledge and process variables, suggesting that highly structured, abstract mathematical knowledge drives the subject's solution strategy. (Author/MK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED183370

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Kathleen Corsale

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Knowledge Base Effects in Children's Number Analogy Solutions[1]

Kathleen Corsale and Drew Gitomer

University of Pittsburgh

[1]Paper presented at the Psychonomic Society Meetings in Phoenix, Arizona, November, 1979.

SE 029 898

Developmental and individual differences in mathematical aptitude were investigated as a function of knowledge structure and processing variables. Results indicated the relative importance of knowledge structure and strategy skills in aptitude test performance. Protocol data further elaborated the interrelationships between knowledge and process variables, suggesting that highly structured, abstract mathematical knowledge drives the subject's solution strategy.

KNOWLEDGE BASE EFFECTS IN CHILDREN'S NUMBER ANALOGY SOLUTIONS

Kathleen Corsale Drew H. Gitomer

University of Pittsburgh

For the most part problem solving research has focused on the study of general solution strategies and process variables. However, several recent investigations demonstrated that knowledge of particular content domains is important in processing information. For example, Chi has demonstrated the importance of knowledge structure in organizing recall of chess information. Likewise, Voss and his colleagues have shown how the differences in the knowledge bases of high and low knowledge subjects can account for differences in the comprehension and recall of domain related material. The extension of this notion in the study of problem solving has been the investigation of strategy differences in the problem solutions between domain experts and domain novices--most notably in the work of Simon and Simon and Larkin in the study of physics problem solving.

The present study is an attempt to demonstrate the differences between high and low ability children in the area of mathematics problem solving. However, going beyond the expert-novice differences that have previously been demonstrated in problem solving work, the present study attempts first to characterize the nature of the differences in the knowledge bases of high and low ability subjects, and then to detect how features of the knowledge representation influence subjects' problem solving solutions. We will present two kinds of data which directly speak to the interactions between knowledge representation and use of problem strategies. An initial set of data is used to characterize the knowledge representations of subjects and the second set consists of protocol data from subjects' problem solutions. We were also interested in whether the nature of these interactions would change developmentally and so investigated the problem at two grade levels.

We studied children's performance on mathematical analogy problems--a problem commonly found on tests of mathematical aptitude-- because it can be segmented to study performance at various stages of solution. Number analogies were used because the mathematical domain is a rich, yet highly structured domain of knowledge whose structure is fairly limited in contrast to other areas of knowledge. Table 1 indicates the type of problem that was used. The problems were of the form A:B as C:D as E:blank. The child's task was to determine a mathematical relation for the AB pair that was identical to the CD relation and to apply that relationship to E in order to obtain the answer F. The children in the study were fifth and eighth graders. Fifth graders have at least learned multiplication concepts and eighth graders have learned more complex concepts such as fractions, exponential relationships, etc., which permitted analyses of problem solving as a function of different ranges of knowledge.

We initially gave fifth and eighth graders a standardized analogy test consisting of 16 multiple choice problems. Within the same week, all children were given two tasks designed to tap the representations of their knowledge of numerical relationships. These tasks were used to detect the salience of mathematical concepts in the knowledge representations of high and low ability subjects. The first task was a grouping task in which the child was given a matrix of numbers from 0-32, asked to select groups of numbers that went together, and to justify his/her groupings. On the basis of these justifications, we were able to classify the kinds of groups children made into four types. Examples of these types are shown in the top half of Table 2. Abstract concepts represented mathematically-based groupings with superordinate labels such as the set of primes, multiplicative or exponential relationships. Operational concepts involved the stringing together of numbers into number sentences. Non-mathematical concepts were idiosyncratic groupings or groups based on orthographic similarities. And finally, digit based groupings in-

volved numbers that shared common digits, the set of single digit numbers, etc.

The second knowledge representation task was called the pairs task and subjects were presented with 20 pairs of numbers and asked to state as many relationships as they could for each pair. We tallied the total number of different relationships generated by each child for each pair type. Pair types are shown in the bottom of Table 2. We also measured the number of relations generated that we judged as critical for solving analogy problems. For example, saying that 28 and 21 were both in the 20's would not be a critical reason in a typical number analogy context whereas "21 is 3/4 of 28" would be.

Because of the large number of measures that were derived from the grouping task and the pairs task, we attempted to reduce the data by means of a factor analysis. All subjects' data from these two tasks resulted in a three factor solution. The factor weights were then used to derive factor scores for each subject. The first factor was readily interpretable as an estimate of the degree of "abstractness" found in the subjects' groupings and pair relationships. The second was a non-mathematical factor seeming to estimate the number of groupings and relations formed. The third factor represented a preference factor in which operational or computation-based groups and relations were preferred or were more salient than abstract groupings.

Each of the factor scores was entered into multiple regression analyses with analogy test performance as the criterion variable. This was done separately at each grade level. The resulting regression equations are indicated in Table 3 of the handout. As can be seen, for both grades, the linear combination of these variables accounted for approximately 50% of the variance. Although knowledge representation and process factors cannot be dichotomized, this level of prediction

in the absence of traditional process measures is, we think, significant. These data reiterate previous demonstrations of the importance of knowledge representation in cognitive processing. For fifth graders, only the abstract knowledge factor and the generativity factor significantly predicted performance on the standardized analogy test whereas for eighth graders, the salience of operational over abstract knowledge was also a significant negative predictor. This is a developmentally important difference which indicates that at earlier age levels, before more abstract mathematical concepts are taught, the salience of abstract concepts does not differentiate high and low ability children and that this salience tradeoff becomes important as children acquire more abstract concepts. The primary conclusion to be drawn from these regressions is that degree of abstractness in mathematical knowledge is an important predictor of success in analogical problem solving.

Having demonstrated that the form of knowledge representation is a critical variable, the interactions of knowledge representation and strategy usage for children of different abilities were examined. At each grade level, the five highest and the five lowest scorers on the standardized analogy test were selected and engaged individually in a session of oral problem solving. The subject was presented with a number analogy problem and asked to solve it aloud. Problems were serially presented such that the subject was first presented with the AB pair and asked to generate analogically plausible rules. He was then presented with both the AB and the CD pairs simultaneously so that he could test his hypothesized relationship and, if necessary, reformulate his rule. He was then presented with E and asked to generate the appropriate F. Having generated an answer or indicating that he couldn't generate one, the subject was given a multiple choice of five alternative answers.

Each subject was given 20 protocol problems. For each protocol we determined three kinds of variables: a) the probability of successfully obtaining the analogical rule after each new pair of numbers had been presented, b) the use of a backward strategy in attempting to solve the problem, and c) the kinds of errors that were made during the course of solution.

The analysis of error data and the backward strategy data indicate that knowledge representation drives the solution strategy by defining the limits of the problem domain. Thus, for example, high ability subjects who have clear, high level mathematics concepts that are abstract in nature, limit their analogical hypotheses to a few plausible mathematical relationships. Low ability children, in contrast, have very diffuse and lower ordered mathematical concepts and their analogical solutions indicate that they do not solve analogies with systematic, mathematically-based rules.

Let's look first at the strategy data. As mentioned earlier, when, during the course of a new pair presentation, subjects referred back to a previous pair, they were credited with using a backward strategy. These frequencies were collapsed over problems and proportion of problems in which the subject went backward was submitted to Chi square analysis. Probability of going backward was high for all subjects, ranging from 50% to 80%. So the use of the backward strategy did not differentiate either grade or ability level. However, we looked only at those protocols in which the subject had not initially gotten the AB relationship and did use a backward strategy when presented with the CD pair.

For just these subjects, the probability of subsequently arriving at the correct analogical rule was examined to determine if such a strategy would increase the likelihood of obtaining the correct rule. The number of subjects obtaining the rule under these conditions is presented in Table 4 of the handout. As can be seen, there is indeed a relationship at both grade

levels between getting the relationship and ability level. On the surface, a relationship between ability level and obtaining the correct analogical rule seems a trivial finding. However, these are all subjects who initially failed to find the correct analogical rule and who engaged in what can be characterized as a backwards strategy. Thus, even though use of a backwards strategy does not differentiate ability level subjects, high ability subjects use the backwards stratgy more effectively: 2/3 of the high aptitude subjects arrive at the correct rule after such a strategy whereas only 1/5 of the low ability subjects do. To demonstrate that the backward strategy is mediating the effective performance of high ability subjects, their probability of obtaining the correct relationship when they do not use a backward strategy is only about 25%.

A look at the error data will help to explain the ability difference in the effectiveness of the backward strategy. First, it must be stated that low ability subjects have many more errors at all phases of problem solution than the high ability subjects. The errors were defined as either violations of mathematical constraints or of analogical contraints. Examples of the different error types are presented in Table 5. Mathematical violations were of two types: computation errors or digit errors in which the subject treated a number not as a total number concept, but as a set of isolated digits. The statement "64 and 16 go together because they both have a 6 in them" is an example of a digit error. Analogical errors were of 6 types. Non-restrictive errors involved obtaining commonalities within pairs, but no relationship that was specific to A and B or C and D. Series errors involved turning an analogy problem into a series problem. Single pair errors occurred when the subject adopted a rule to apply to E that was true only of AB or of CD, but not both. Subjects who looked for relationships across pairs rather than within pairs were considered to have made an ACBD violation. Non analogical computation violations involve computations that were analogically inappro-

priate. Finally, direction violations were those in which the subject applied the correct rule but in the wrong direction.

Table 6 indicates the mean errors across all problems and subjects for each grade and ability group. The most obvious result is the lack of grade differences in error frequencies. Rather, any differences in error frequency seem to be due to ability differences. This lack of a grade effect was recurrent in all the analyses and will be discussed later.

Analysis of variance indicated a three way interaction of grade, ability level and error type. Low ability children at both grade levels commit more non-analogic computation, non-restrictive and digit errors. They also make more mathematical computation errors although this difference is not significant with an n of 5. The kinds of errors they make indicate that they do not restrict their hypotheses concerning an analogical rule to mathematical concepts (as noted by the digit errors) or to analogical concepts (as noted by non-analogic computation). These findings substantiate the data presented earlier which indicated that these low ability subjects have more diffuse, less-structured knowledge representations of mathmatics. Here we see the manifestation of their knowledge representation in the kinds of errors they make.

The parallel between knowledge representation and solution strategy can also be seen in errors of the high ability subjects. Although the series error is not significant in differentiating between high and low ability subjects, it is the only error which is committed more frequently by high than by low ability subjects. However, this is a sophisticated kind of error involving the detection of mathematical relationships across pairs that follows a constrained rule. Again, the knowledge representation data which indicated tightly constrained math concepts for these subjects parallels their use of that knowledge as seen in the kinds of errors they make.

With regard to the low frequency of errors for the high ability children, it is the case that when they could not detect a rule, high ability children would "give up" and not select a multiple choice answer whereas low ability children would select an answer--usually a wrong one--and justify it post-hoc. We looked only at children who still had not obtained the correct analogical rule when preserted with the multiple choices. The frequencies of subjects who gave up in each condition were presented in Table 7. They indicate that high ability children operate within both mathematical and analogical constraints in order to achieve a goal of an analogically correct answer. They know when they are wrong and give up rather than choose an answer that they know is wrong. Not only do low ability children choose the wrong answer rather than give up, anecdotal evidence suggests that they are perfectly happy with and can justify their choices on non-analogical and/or non-mathematical grounds.

Considering both the knowledge representation data and the protocol data, the findings suggest first, that high ability subjects have a greater degree of abstract mathematical knowledge and a greater salience of abstract over operational concepts. Second, this knowledge correlates with and predicts analogy performance. Finally, and most importantly, high ability subjects use their knowledge of abstract number relationships to constrain the domain of permissible operations. That is, the knowledge representation sets conditions on the appropriate use of strategies.

As an illustration of this last point, consider the differences between high and low ability fifth graders who were not very familiar with exponential relationships. High ability children were able to limit their hypotheses to multiplicative ones and often came up with the correct exponential rule by limiting their operations to mathematical ones and by looking for relations between AB and CD. Low ability children, on the other hand,

often engaged in non-analogical computation and non-restrictive errors. They did not have the highly constrained organizational structure of abstract knowledg that would provide them with constrained rules of operation.

Finally, the virtual absence of grade effects in these data would suggest that individual differences in ability are an important source of variance in these tasks. Indeed, the ability differences suggest that, at least in some areas, individual differences would be a better developmental variable than either age or grade level.

TABLE 1

Problem	Rule
7:21 :: 5:15 :: 4:_	Simple multiplicative relationship (x3)
15:19 :: 8:12 :: 9:_	Simple additive relationship (+4)
10:40 :: 6:36 :: 5:_	Tricky additive relationship (+30)
28:21 :: 24:18 :: 20:_	Proportion relationship (x 3/4)
6:0 :: 10:2 :: 18:_	Two-operation relationship (+2, -3)
1000:10 :: 64:4 :: 125:_	Exponential relationship ($\sqrt[3]{}$)

- -

TABLE 2

Knowledge Base Tasks

I. Grouping Task

Abstract Groupings	Operational Groupings	Digit-based Groupings	Non-Mathematical Groupings
Multiples	Number sentences	Single digit #s	Idiosyncratic
Exponents	Computation	Decade subsets	No Reason
Primes		Common digit	Orthographic
Composite			Proximity
Factors			
Odd/Even			

II. Pairs Task

Pair Type	Total # of Relationships Generated	# of critical Relationships Generated
Additive (15, 19)		
Multiplicative (7, 28)		
Factor (18, 24)		
Exponential (7, 49)		

- -

TABLE 3

Regression Analyses

Criterion variable = performance on a standardized number analogy test

5th graders: 1.26(abstract K) + .67 (generativity)*

8th graders: 1.34(abstract K) + .41 (generativity) - .41 (operational - abstract) **

*r^2 = .51
**r^2 = .46

TABLE 4

Relationship between obtaining the analogical rule and ability level for subjects who fail to identify the A:B relationship and who use a backward strategy when presented with the C:D pair.

5th graders:

		Ability	
		low	high
correct relationship	yes	5	22
	no	31	16

$\chi^2 = 15.45$, $p \leq .001$

8th graders:

		Ability	
		low	high
correct relationship	yes	16	23
	no	59	11

$\chi^2 = 21.84$, $p \leq .001$

TABLE 5

Error Types and Examples

I. Mathematical
- A. Computation - arithmetic error such as"12 x 13 is 169"
- B. Digit - treating numbers as a set of digits such as "64 and 16 are related because they both have a 6 in them"

II. Analogical
- A. Non-restrictive (NR) - a correct characterization of number pairs that is not specific enough to allow an analogical solution such as "11 and 33 go together because they are both odd"
- B. Series (Ser) - stating a serial pattern across pairs. For example, in the problem 28:21 :: 24:18 :: 20:_, the subject may say that the difference between the first and second numbers decreases by 1 with each successive pair.
- C. Single Pair (SP) - using a correct relationship that applies only to the A:B or the C:D pair, but not both. For example, in the problem 9:18 :: 6:15 :: 3:_, where the rule is actually +9, the subject may apply a x2 relationship derived from consideration of only the first pair.
- D. AC-BD (AC-BD) - looking for relationships across pairs instead of within pairs. For example, 10:40 :: 6:36 :: 5:_, the subject may notice that 10 - 6 = 4 and that 40 - 36 = 4, yet be unable to come up with the correct analogical rule of A + 30 = B.
- E. Non-analogical computation (NAC) - computations which are not constrained by identical pairwise relationships. For example, 2:4 :: 8:64 :: 5:_, the subject may say that the answer is 16 because 2 x 8 is 16.
- F. Directional (DIR) - applying the correct operator, but in the wrong direction. For example, if the rule is to multiply by 3, the subject may divide by 3.

TABLE 6

Frequencies of Errors by Grade and Ability

Grade	Ability	MC	MD	NR	Ser	SP	AC-BD	NAC	Dir
5	Low	6.0	9.4	6.8	1.0	3.4	0.2	14.2	0.4
	High	1.8	2.8	1.6	3.2	0.8	0.8	0.4	0.6
8	Low	6.4	3.6	15.6	1.2	3.0	0.0	8.4	1.6
	High	0.6	2.0	1.4	3.0	0.2	1.4	0.2	0.0

TABLE 7

Relationship between giving up on a problem and ability level for subjects who have not obtained the correct analogical rule when multiple choice alternatives are presented

5th graders:

		Ability	
		Low	High
Give up	yes	24	16
	no	33	2

$\chi^2 = 12.03$, $p < .001$

8th graders:

		Ability	
		Low	High
Give up	yes	9	8
	no	35	10

$\chi^2 = 3.69$, $p < .055$